IN THE SPECIFICATION:

Please amend the paragraph starting at page 1, line 24 and ending at page 2, line 4 as follows.

--In the browser in the graphical user interface, movement to next contents and input in a form are performed through mouse operation and keyboard ~~entrance~~ entry, but in the voice browser, they are done through voice input. That is, a user's voice input is voice-recognized, and the recognition result is used to perform movement to next contents and input in the form.--

A1

Please amend the paragraph starting at page 2, line 5 and ending at line 14 as follows.

--There is a method in which a dedicated markup language is used as these contents for voice browsers. In this method, however, access cannot be made to the contents by the browser of the graphical user interface, and with this voice browser, access cannot be made to contents for the graphical user interface that currently exist numerously. Thus, there is a method in which ~~HTML;~~ HTML, a markup language that is used in the browser of the graphical user ~~interface~~ interface, is used also in the voice browser.--

A2

Please amend the paragraph starting at page 2, line 15 and ending at line 20 as follows.

--In this method, output contents and input candidates in voice, namely contents of processing suitable for voice recognition vocabularies and ~~man-power~~ man-power, are

A3

determined from contents written in HTML, according to a specific rule. For example, there is a voice browser apparatus using rules as described below.

Please amend the paragraph starting at page 2, line 21 and ending at page 3, line 6 as follows.

--First, output contents shall constitute the text ranging from the head to the end of the HTML document to be subjected to browsing. However, if the URL indicates some midpoint in the HTML document, the output contents shall cover the range therefrom, and if there is ~~a~~ an <HR> tag at some midpoint, the output contents shall cover the range ending with the tag. The input candidate shall constitute an anchor in the same range (text in the range surrounded by the <A> tag). When a word existing in the input candidate is inputted, the target to which it is linked is defined as a new object of browsing to perform similar processing.

Please amend the paragraph starting at page 3, line 7 and ending at line 21 as follows.

--For example, the case where the HTML document shown in FIG. 4 is targeted will be discussed. Assume that the URL of this HTML document is "http://guide/index.html". First, the voice browser outputs "Please select a genre of shops from the following. French. Italian." with a voice, and waits for a user's input. When the user inputs "Italian" with a voice, for example, the voice browser performs similar processing from the position of the HTML document of "http://guide/index.html # italian". In other words, it outputs "Please select a ~~shops~~ shop. ▽▽. □□.", and waits for the user's input. When the user inputs "▽▽", for example, it obtains the HTML document of "http://guide/shop3.html" to carry out similar processing.--

Please amend the paragraph starting at page 4, line 2 and ending at line 9 as follows.

--Thus, an objective of the present invention is to provide a voice browser apparatus in which a plurality of rules for defining output contents and input candidates in the form of voice from contents written in markup language for the graphical user interface, such as ~~HTML~~ HTML, is prepared, thus allowing a user or a content creator to designate which rule of them is used.--

Please amend the paragraphs starting at page 4, line 10 and ending at page 7, line 14 as follows.

--According to ~~one aspect,~~ an aspect of the present invention ~~which achieves the objective relates to a document processing apparatus comprising document obtaining means for obtaining a document written in predetermined markup language from a designated source from which the document is to be obtained~~, rule selecting means selects ~~for selecting~~ a rule defining voice input/output contents from a plurality of predetermined rules, document analyzing means ~~for analyzing~~ analyze a designated range of ~~the~~ a document obtained ~~by the above described document obtaining means~~, based on the rule selected by the ~~above described~~ rule selecting means, ~~to fetch~~ and voice output contents, voice input candidates, and designation information are fetched. ~~for designating a next object of processing corresponding to each voice input candidate, voice outputting means for voice-outputting the voice output contents fetched by the above described document analyzing means, voice recognizing means for voice-recognizing the voice input from the user, and controlling means for checking the result of recognition by the above described voice recognizing means against the input candidates fetched by the above~~

described document analyzing means to control obtainment of a new document by the above described document obtaining means or next analysis by the above described document analyzing means, based on designation information corresponding to the input candidate matching the recognition result.

According to another aspect, the present invention which achieves these objectives relates to a document processing method comprising a document obtaining step of obtaining a document written in predetermined markup language from a designated source from which the document is to be obtained, a rule selecting step of selecting a rule defining voice input/output contents from a plurality of predetermined rules, a document analyzing step of analyzing a designated range of the document obtained in the above described document obtaining step, based on the rule selected in the above described rule selecting step, to fetch voice output contents, voice input candidates, and designation information for designating a next object of processing corresponding to each voice input candidate, a voice outputting step of voice-outputting the voice output contents fetched in the above described document analyzing step, a voice recognizing step of voice-recognizing the voice input from the user, and a controlling step of checking the result of recognition by the above described voice recognizing step against the input candidates fetched in the above described document analyzing step to control obtainment of a new document by the above described document obtaining step or next analysis by the above described document analyzing step, based on designation information corresponding to the input candidate matching the recognition result.

According to still another aspect, the present invention which achieves these objectives relates to a computer-executable program for controlling a computer to perform document processing, said program comprising codes for causing the computer to perform a

~~document obtaining step of obtaining a document written in predetermined markup language from a designated source from which the document is to be obtained, a rule selecting step of selecting a rule defining voice input/output contents from a plurality of predetermined rules, a document analyzing step of analyzing a designated range of the document obtained in the above described document obtaining step, based on the rule selected in the above described rule selecting step, to fetch voice output contents, voice input candidates, and designation information for designating a next object of processing corresponding to each voice input candidate, a voice outputting step of voice-outputting the voice output contents fetched in the above described document analyzing step, a voice recognizing step of voice-recognizing the voice input from the user, and a controlling step of checking the result of recognition by the above described voice recognizing step against the input candidates fetched in the above described document analyzing step to control obtainment of a new document by the above described document obtaining step or next analysis by the above described document analyzing step, based on designation information corresponding to the input candidate matching the recognition result.~~--

Please amend the paragraph starting at page 10, line 23 and ending at page 11, line 3 as follows.

--A speaker 204 outputs voice data generated by the voice output portion 107. A microphone 205 inputs voice data that is processed by the voice input portion 108. A network interface 206 achieves communication via a network at the time when the HTML document obtaining portion 101 obtains the HTML document through the network. A bus 207 connects the above described ~~each portion~~ portions.--

Please amend the paragraph starting at page 11, line 4 and ending at line 6 as follows.

--~~Processing~~ A processing procedure of the voice browser apparatus of this embodiment will be described below, referring to the flowchart in FIG. 3.--

Please amend the paragraph starting at page 12, line 10 and ending at page 13, line 1 as follows.

--The rule used in this embodiment will now be described. In this embodiment, the rule in the case where the value for the designation rule storing portion 104 is "H" is as follows. Initial output contents shall be the value of the OUTPUT attribute of the <VB> tag and input candidates that will be described subsequently. The input candidates shall be respective indexes surrounded by the <H> tag in the HTML document. When a statement included in the input candidate is inputted, the following processing is performed. First, next output contents shall constitute the text ranging from the selected index to the next <H> tag or to the end of the document. And the input candidate shall constitute ~~a~~ an anchor in the same range (text in the range surrounded by the <A> tag). When a statement included in the input candidate is inputted, the target to which it is linked is defined as a new object of browsing to perform similar processing.--

Please amend the paragraph starting at page 13, line 2 and ending at line 16 as follows.

--On the other hand, in this embodiment, the rule in the case where the value for the designation rule storing portion 104 is "L" is a rule to perform the processing procedure

described as a prior art. That is, output contents shall be the text ranging from the head to the end of the HTML document that is an object of browsing. However, if the URL indicates some midpoint in the HTML document, the output contents shall cover the range therefrom, and if there is a an <HR> tag at some midpoint, the output contents shall cover the range ending with the tag. The input candidate shall constitute an anchor in the same range. When a statement included in the input candidate is inputted, the target to which it is linked is defined as a new object of browsing to perform similar processing.--

Please amend the paragraph starting at page 13, line 17 and ending at line 23 as follows.

--In Step S303, in accordance with the rule appropriate of to the value stored in the designation rule storing portion 104, the HTML document stored in the HTML document storing portion 102 is analyzed to fetch the contents of voice input/output and stores the same in the input/output storing portion 106. Then, a movement to Step S304 is made.--

Please amend the paragraph starting at page 14, line 14 and ending at line 23 as follows.

--In the former case, the value of the OUTPUT attribute of the <VB> tag, and the input candidate that will be described subsequently is are stored in the area 601 of the input/output contents storing portion 106. Also, each index surrounded by the <H> tag in the HTML document is stored in the column 603 as the input candidate. And, the URL of the HTML document currently under processing is stored in the column 604 for each index. In addition, the pattern including the tag of each index is stored in the column 605.--

Please amend the paragraph starting at page 15, line 8 and ending at line 20 as follows.

--On the other hand, if the value stored in the designation rule storing portion 104 is "L", text ranging from the head to the end of the HTML document is stored in the area 601 as voice output contents. However, if the URL indicates some midpoint of the HTML document, the range shall start therefrom, and if there is ~~a~~ an <HR> tag at some midpoint, the range shall end with the tag. Then, the input candidate is defined as the anchor in the same range, and the URL of the target to which it is linked is stored in the column 604 for each candidate. The column 605 shall be empty. FIG. 6 shows a state of the input/output contents storing portion 106 when the HTML shown in FIG. 5 is processed.--

A14

Please amend the paragraph starting at page 16, line 6 and ending at line 12 as follows.

--In Step S306, the result of the voice recognition in ~~step~~ Step S305 is compared with the input candidates stored in the input/output contents storing portion 106. If there is an input candidate matching the result, a movement to Step S307 is made. If there is no candidate matching the result, a return to ~~step~~ Step S305 is made.--

A15

Please amend the paragraph starting at page 16, line 20 and ending at line 24 as follows.

--In Step S308, an HTML document shown by the URL of the input candidate for which matching has been obtained in ~~step~~ Step S306 is newly obtained and is stored in the HTML document storing portion 102. Then, a return to Step S302 is made.

A16

Please amend the paragraph starting at page 16, line 25 and ending at page 17, line 5 as follows.

--The HTML document of FIG. 5 is stored in the HTML document storing portion 102, and if "Italian" is inputted when the input/output contents storing portion 106 is in the ~~sate~~ state shown in FIG. 6, the input/output content storing portion 106 newly enters a state as shown in FIG. 7. Thus, the input/output after the HTML document in FIG. 5 is stored in the HTML document storing portion 102 is as follows.--

Please amend the paragraph starting at page 18, line 9 and ending at line 14 as follows.

--The HTML document in FIG. 10 is different from the HTML document in FIG. 5 only in the value of the MODE attribute of the <VB> tag. Use of the voice browser apparatus of this embodiment makes it possible to change the contents of voice interaction for the similar HTML document by only ~~by~~ changing part of the tag.--

Please amend the paragraph starting at page 20, line 18 and ending at page 21, line 3 as follows.

--In the above described embodiments, the case where the rule directly designated by the user is defined as a user rule has been described, but the present invention is not limited thereto, and it is also possible to store in advance the rule to be applied for each HTML document and apply the stored rule each time the HTML document is processed. This can be achieved by storing in advance a table in which the URL of the HTML document ~~is corresponded~~ corresponds to the rule to be applied, using the URL to search the table each time the HTML

document is obtained, and having the corresponding rule stored in the user rule storing portion 1101 if such a URL is stored in the table.--

Please amend the paragraph starting at page 21, line 27 and ending at page 22, line 6 as follows.

--In the above described embodiments, the case where every input/output for the voice browser apparatus is performed using voice has been described, but the present invention is not limited thereto, and inputting means other than voice may be used in part. For example, the number of the input candidate may be inputted with key strokes instead of voice-inputting the input candidate.--

Please amend the paragraph starting at page 24, line 20 and ending at line 25 as follows.

--In the above described embodiments, the case where a program required for operations is stored in the disk device has been described, but the present invention is not limited thereto, and it may be achieved using any storage medium. Also, it may be achieved using a circuit operating in a similar way.--

Please amend the paragraph starting at page 25, line 7 and ending at line 12 as follows.

--Furthermore, as long as the feature of the above described embodiments can be achieved, the present invention may be applied to a system ~~comprised of~~ comprising a plurality

of apparatuses (a computer main body, an interface apparatus, a display, etc.), or may be applied to equipment ~~comprised of~~ comprising a single apparatus.--